The Volatility of the Instantaneous Spot Interest Rate Implied by Arbitrage Pricing - A Dynamic Bayesian Approach

Ramaprasad Bhar* Carl Chiarella † Hing Hung‡
Wolfgang J. Runggaldier§

June 23, 2005

Abstract

This paper considers the estimation of the volatility of the instantaneous short interest rate from a new perspective. Rather than using discretely compounded market rates as a proxy for the instantaneous short rate of interest, we derive a relationship between observed LIBOR rates and certain unobserved instantaneous forward rates. We determine the stochastic dynamics for these rates under the risk-neutral measure and propose a filtering estimation algorithm for a time-discretised version of the resulting interest rate dynamics based on dynamic Bayesian updating in order to estimate the volatility function. Our time discretisation can be justified by the fact that data are observed discretely in time. The method is applied to US Treasury rates of various maturities to compute a (posterior) distribution for the parameters of the volatility specification.

Acknowledgment: The authors wish to thank the referees for helpful suggestions that have improved this paper. The usual caveat applies. Bhar and Chiarella acknowledge financial support from Australian Research Council Grant A00103864.

* School of Banking and Finance, The University of New South Wales, Sydney, r.bhar@unsw.edu.au

† School of Finance and Economics, University of Technology, Sydney, Australia, E-mail: Carl.Chiarella@uts.edu.au (Corresponding author)

‡ School of Finance and Economics, University of Technology, Sydney, Australia, E-mail: hing.hung@uts.edu.au

§ Dipartimento di Matematica Pura ed Applicata, Università di Padova, E-mail:runggal@math.unipd.it

1 Introduction

The literature on the estimation of spot interest rate models in particular and term structure dynamics in general continues to burgeon. Perhaps the most influential early work has been that of Chan et al. (1992) (henceforth CKLS). Their estimation of the constant elasticity of the diffusion term with respect to the instantaneous spot rate has been repeated for many different markets, time periods and estimation procedures. CKLS used the generalised method of moments whilst Nowman (1997, 1998, 2001, 2003) applied Gaussian estimation techniques and Babbs and Nowman (1999) used Kalman filtering methods. Using Gaussian estimation techniques, Episcopos (2000) estimated the parameters of the CKLS type specification for a number of markets. He obtained estimates of the spot rate elasticity of the diffusion term much lower than that obtained by CKLS. Sun (2003) considered a nonlinear diffusion term and also allowed for GARCH effects, but also found lower values for the elasticity in several markets. All of these estimation methodologies yield point estimates and a feature of these various empirical studies is the wide range of point estimates obtained.

Another issue related to the estimation of the spot interest rate process is what data should be used to proxy the unobserved instantaneous spot rate. Typically most studies have used US one-month Treasury bill rates (CKLS and Nowman (1997)), US three-month Treasury bill rates (Sun 2003), one month Euro-currency rates (Nowman 1998) and one-month interbank rates (Sun (2003) and Episcopos (2000)). The results of Chapman et al. (1999) suggest that the choice of proxy variable (in particular whether it be a one-month or three-month rate) should not lead to a great deal of error in the estimation procedure. However, given that it is not difficult to obtain the dynamics for, say, discrete tenor three-month rates implied by a particular instantaneous spot rate process, it seems strange that the literature has not developed in the direction of estimating directly the parameters of the processes for observed market rates. This will be one of the contributions of the current paper.

In this paper we use the framework of Heath, Jarrow and Morton (1992) (henceforth HJM) to model the dynamics of the interest rate market. The starting point of HJM is a specification of the dynamics of the forward rate to any general maturity. We specify a forward rate volatility function that yields the same volatility function for the instantaneous spot rate of interest considered in the earlier cited literature. An important difference is that the dynamics of the interest rate processes occur under the risk-neutral measure. Under this measure the HJM procedures enable us to obtain the dynamics of pure discount bond prices. These can in turn be related to the discretely compounded LIBOR rates. This link then enables us to determine the dynamics for LIBOR rates. It turns out that the dynamics of the LIBOR rate, the instantaneous spot rate of interest and another instantaneous forward rate evolve simultaneously under the risk-neutral measure.

The link between pure discount bond prices and LIBOR rates mean that these rates can be regarded as observable under the risk-neutral measure, whilst the other two instantaneous rates referred to in the previous paragraph are not observable. We are thus dealing with a partially observed stochastic dynamic system whose estimation may be undertaken by the use of nonlinear filtering methods. Here we develop a dynamic Bayesian updating algorithm analogous to the one proposed in Chiarella, Pasquali and Ruggaldier (2001). The basic approach proposed here has been applied to a much simpler (and approximate) representation of discrete tenor interest rate dynamics in Bhar, Chiarella and Runggaldier (2002). The potential of Bayesian methods in the estimation of financial models is starting to be appreciated, see e.g. Polson and Tew (2000). A range of implementations are possible, the one presented here has been chosen because we are able to rigorously demonstrate its convergence properties.

The paper makes three main contributions. First, the specification of the interest rate dynamics allows us to use as observations interest rates of any maturity. In particular we use interest rates of much longer maturity (6 and 12 months) than those usually used in the literature on estimation of interest rate models. Second, we demonstrate the feasibility of the Bayesian updating filtering algorithm as a tool for estimating interest rate models within the HJM framework. Third, we compute a posterior distribution for the parameter values, rather than just the point estimates of the traditional literature. This gives a better understanding for the range of point estimates obtained in the literature.

The plan of the paper is as follows: in section 2 we derive the stochastic dynamic system followed by the instantaneous spot rate and discretely compounded LIBOR rates. Since the data are observed in discrete time, in section 3 we outline the way in which the continuous time stochastic differential equation system is discretised. In section 4 we outline the way in which the dynamic Bayesian updating algorithm is applied to the estimation problem. In section 5 we discuss implementation issues and apply the algorithm to some U.S. data. Section 6 concludes and makes suggestions for future research. Detailed technical derivations are relegated to the appendices.

2 The Dynamics of LIBOR Rates Implied by HJM Bond Prices

We use the Brace and Musiela (1994) (henceforth BM) parameterisation of the HJM model, which is in terms of $r(t,x)(x \geq 0)$ the x-period instantaneous forward rate at time t for maturity $t+x$ (see figure 1 [1]). Under the risk-neutral measure $\widetilde{\mathbb{P}}$ this rate satisfies the stochastic integral equation

$$r(t,x) = r(0,t+x) + \int_0^t \sigma(s,t+x)\overline{\sigma}(s,t+x)ds + \int_0^t \sigma(s,t+x)d\widetilde{w}(s), \tag{1}$$

where $r(0,x)$ is the initial forward curve, $\widetilde{w}$ is a Wiener process under $\widetilde{\mathbb{P}}$ and $\sigma(t,x)$ is the instantaneous forward rate volatility function that could (and in our application will) depend on certain instantaneous forward rates. In equation (1)

$$\overline{\sigma}(s,t+x) = \int_s^{t+x} \sigma(s,u)du. \tag{2}$$

It is important to stress that even though we use the BM parameterisation for the forward rate dynamics, we use the notation for the volatility function as in HJM in that $\sigma(t,x)$ refers to

[1] In figures 1 and 2 we show the typical investor standing at time 0 to represent the fact that all dynamics are perceived from the perspective of time 0 where the relevant information is the currently observed forward curve $r(0,t)$.

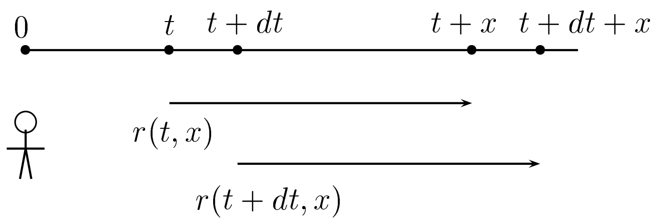

Figure 1: Time line for the BM forward rate.

the forward rate volatility at time t applicable for time $x(\geqslant t)$. This is in contrast to BM who use the volatility function $\tau(t,x)$ to denote the forward rate volatility at time t applicable for time $t+x$. Of course these two different specifications of the forward rate volatility function are related via

$$\tau(t,x) = \sigma(t,t+x)$$

and one may work with either specification. For our application it turns out to be more convenient to use $\sigma(t,x)$.

In this notation the instantaneous spot rate of interest $r(t)$ is given by

$$r(t) = r(t,0), \tag{3}$$

and satisfies the stochastic integral equation

$$r(t) = r(0,t) + \int_o^t \sigma(v,t)\overline{\sigma}(v,t)dv + \int_0^t \sigma(v,t)d\widetilde{w}(v). \tag{4}$$

The price at time t of a $(t+x)$- maturity zero coupon bond is related to $r(t,x)$ by

$$b(t,x) = exp\left(-\int_0^x r(t,u)du\right). \tag{5}$$

Next we relate the x-period LIBOR rate to the bond price $b(t,x)$. We then derive the relationship between the bond price and the underlying state variables (a set of discrete tenor forward rates) upon which the forward rate volatility function depends. The dynamics of these state variables determine the evolution of the forward curve.

Consider a time period $(0,t)$ over which we have a set of observations of the x-period LIBOR rate, that we denote $L_x(t)$.This is an annualised rate at which $\$1$ invested at time t compounds simply to become $\$(1+xL_x(t))$ at time ($t+x$).

The LIBOR rate $L_x(t)$ is related to the continuously compounded Brace-Musiela instantaneous forward rate by (see figure 2)

$$1 + xL_x(t) = exp(\int_o^x r(t,u)du). \tag{6}$$

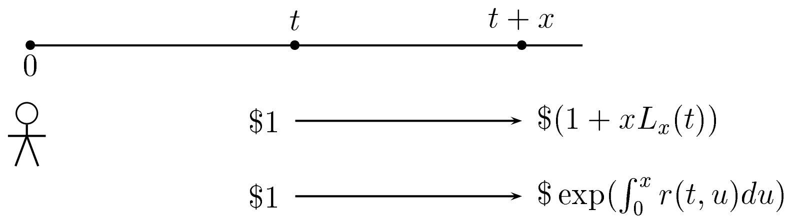

Figure 2: The LIBOR rate $L_x(t)$

From equations (5) and (6) we deduce the relationship between the LIBOR rate and the bond price, viz

$$L_x(t) = \frac{1}{x}\left(\frac{1}{b(t,x)} - 1\right). \tag{7}$$

However it turns out to be more convenient to work in terms of the quantity

$$l_x(t) = L_x(t) + \frac{1}{x}, \tag{8}$$

which is related to $b(t,x)$ via

$$l_x(t) = \frac{1}{xb(t,x)}. \tag{9}$$

We consider the volatility function of the general form[2]

$$\sigma(t,u) = g(r(t,x_1), \cdots, r(t,x_n))e^{-\lambda(u-t)}, \tag{10}$$

where $\vec{r}(t,\cdot) \equiv [r(t,x_1), \cdots, r(t,x_n)]$ is a vector of discrete tenor forward rates chosen in the belief that these particular maturities most affect the evolution of the forward curve e.g. perhaps they correspond to the most liquid maturities. In our subsequent application we shall specialise (10) to the case where $g(\cdot)$ depends on just one argument and has the particular form

$$g(r) = \sigma_0(\min\{\varepsilon^{-1}, \max[|r|, \varepsilon]\})^\delta \tag{11}$$

where $\delta > 0$ and σ_0 are parameters to be estimated and $\varepsilon > 0$ is a given, arbitrarily small constant.

This representation is consistent with the earlier cited empirical literature that concentrates merely on dependence on the instantaneous short rate. We use $\min\{\varepsilon^{-1}, \max[|r|, \varepsilon]\}$ to the prevent the volatility from becoming either zero or infinite.

Thus equation (10) specialises to

$$\sigma(t,u) = g(r(t))e^{-\lambda(u-t)}. \tag{12}$$

[2]We write a minus sign in front of the λ since we would normally expect λ to be positive reflecting the empirically observed fact that usually interest rate volatility decreases as maturity increases, though this relation often inverts at very short maturities.

Subsequent applications could allow for dependence on a number of discrete tenor forward rates.

Chiarella and Kwon (2003) show that with the specification (12) the bond price may be expressed as a deterministic combination of two discrete tenor forward rates $r(t, x_1)$, $r(t, x_2)$ whose tenors may be chosen arbitrarily. The relevant details are summarized in Appendix 2 from equation (73) of which we have

$$b(t,x) = exp(-[\bar{b}_o(t,x) + \sum_{i=1}^{2} \bar{b}_i(t,x) r(t,x_i)]), \tag{13}$$

where the $\bar{b}_i(t,x)(i=0,1,2)$ are defined in general by equations (71) and (74) and evaluated for the specific volatility function (12) in equations (85)-(87). The stochastic differential equations followed by the $r(t,x_k)(k=1,2,\cdots,n)$ are given by equations (76) of appendix 1, namely [3]

$$\begin{aligned} dr(t,x_k) = [b_0^{'}(t,x_k) + \sum_{i=1}^{2} b_i^{'}(t,x_k) r(t,x_i) + \sigma(t,t+x_k)\bar{\sigma}(t,t+x_k)]dt \\ + e^{-\lambda x_k} g(r(t,x_1),...,r(t,x_n)) d\widetilde{w}(t), \qquad (k=1,2,...,n). \end{aligned} \tag{14}$$

Keeping in mind that our aim is to estimate the parameters $(\sigma_0, \delta, \lambda)$ used to specify the particular volatility structure (12), we use the foregoing term structure dynamics as follows. First we treat equation (9) for $l_x(t)$ (with $b(t,x)$ given by (13)) as the observation equation, with underlying unobserved state variables $r(t,x_1), r(t,x_2), \cdots, r(t,x_n)$ being driven by the system (14). Note that here we have set things up in such a way that the $r(t,x_1), r(t,x_2)$ appearing in (13) are the first two elements of the vector $\vec{r}(t,\cdot)$ upon which the volatility function is dependent. It should be stressed that this choice is somewhat arbitrary and any two elements of $\vec{r}(t,\cdot)$ might have been chosen. Indeed it is possible to use two discrete tenor forward rates not belonging to $\vec{r}(t,\cdot)$, in which case an additional two stochastic differential equations for their dynamics would have to be appended to the system (14). The particular choices made in this regard are implementation issues.

Turning to our particular implementation with the volatility function (12), this fits into the general structure of equation (10) by setting $n = 1$ and $x_1 = 0$ so that

$$\vec{r}(t,\cdot) = r(t,0) = r(t).$$

In equation (13) we also set $r(t,x_1) = r(t,0) = r(t)$ and leave $r(t,x_2)$ as some arbitrary tenor discrete forward rate. The dynamics for $r(t,x_2)$ will append an additional stochastic differential equation to the one to which the system (14) reduces in this case.

To summarise, the expression for $b(t,x)$ will be given by

$$b(t,x) = exp(-\left[\bar{b}_0(t,x) + \bar{b}_1(t,x) r(t) + \bar{b}_2(t,x) r(t,x_2)\right]), \tag{15}$$

where the dynamics for $r(t)$ and $r(t,x_2)$ are given by

$$\begin{aligned} dr(t) =& [b_0'(t,0) + b_1'(t,0) r(t) + b_2'(t,0) r(t,x_2)]dt \\ &+ g(r(t)) d\widetilde{w}(t), \end{aligned}$$

[3]Note that $b_i'(t,x) \equiv \frac{\partial}{\partial x} b_i(t,x)$ and the precise expressions are given in equations (82)-(84).

and

$$\begin{aligned}dr(t,x_2) =& [b'_0(t,x_2) + b'_1(t,x_2)r(t) + b'_2(t,x_2)r(t,x_2)\\ &+\sigma(t,t+x_2)\bar{\sigma}(t,t+x_2)]\,dt\\ &+ e^{-\lambda x_2}g(r(t))d\widetilde{w}(t).\end{aligned}$$

As we have stated, the choice of x_2 is arbitrary, for an initial implementation we choose x_2 to be the same as the tenor x of the observed LIBOR rates.

Here we should stress that the driving dynamics (14) are under the risk neutral measure $\widetilde{\mathbb{P}}$. However the LIBOR rates are observed under the real world measure $\mathbb{P}$. To convert the dynamic (14) to the dynamics under $\mathbb{P}$ we would have to introduce the market price of interest rate risk. However the diffusion of the underlying process will be the same under $\mathbb{P}$ and $\widetilde{\mathbb{P}}$. Of course the drifts will differ under the two measures, but we are not concerned in this paper with estimating the drift term rather we focus just on estimating the volatility function. If we were interested in estimating the drift as well then we would have to make some assumption about the market price of risk; for instance, either it depends on the state variables in some way or it follows some stochastic process. Then we would need to consider the dynamics and estimation procedure under the historical measure. The Bayesian updating algorithm (appropriately modified) to be described below could still be applied to the resulting stochastic dynamical system.

3 State Space Form of the Model

Summarizing the results of the previous section, we shall take as our partially observable system the (unobservable) instantaneous x-period forward rate $r(t,x)$, and instantaneous spot rate $r(t,0) \equiv r(t)$. The set of stochastic differential equations for the state may be succinctly written

$$dr(t,x) = R_x(r(t,x),r(t,0))dt + e^{-\lambda x}g(r(t,0))d\widetilde{w}, \tag{16}$$
$$dr(t,0) = R(r(t,x),r(t,0))dt + g(r(t,0))d\widetilde{w}, \tag{17}$$

where we set

$$\begin{aligned}R_x(r(t,x),r(t,0)) =& b'_0(t,x) + b'_1(t,x)r(t,0) + b'_2(t,x)r(t,x)\\ &+ \sigma(t,t+x)\bar{\sigma}(t,t+x),\end{aligned} \tag{18}$$

and

$$R(r(t,x),r(t,0)) = b'_0(t,0) + b'_1(t,0)r(t,0) + b'_2(t,0)r(t,x). \tag{19}$$

Equations (16), (17) are the state transition equations for the unobserved state variables, $r(t,x)$, $r(t,0)$.

We shall use $S(t) := (r(t,0), r(t,x))^T$ to denote the state vector at time t. This enables us to write the state transition equations (16), (17) as

$$dS(t) = F(S(t);\theta)dt + V(S(t);\theta)d\widetilde{w}, \tag{20}$$

where

$$
\begin{aligned}
F(S;\theta) &= (R(r(t,x),r(t,0)), R_x(r(t,x),r(t,0)))^T, & (21)\\
V(S;\theta) &= (g(r(t,0)), e^{-\lambda x} g(r(t,0)))^T, & (22)
\end{aligned}
$$

and θ is the (to be estimated) parameter vector

$$\theta = (\sigma_0,\ \delta,\ \lambda).$$

Financial implementations of estimation methodologies are usually carried out in a discrete time setting as data are observed discretely. Thus we discretise (16), (17) using the Euler-Maruyama scheme to obtain

$$
\begin{aligned}
r(k+1,x) - r(k,x) &= R_x(r(k,x),r(k,0))\Delta t + e^{-\lambda x} g(r(k,0))\Delta W_k,\\
r(k+1,0) - r(k,0) &= R(r(k,x),r(k,0))\Delta t + g(r(k,0))\Delta W_k, \qquad (23)
\end{aligned}
$$

where, Δt denotes the time step and (see (18), (19)),

$$
\begin{aligned}
R_x(r(k,x),r(k,0)) &= b_0'(k\Delta t,x) + b_1'(k\Delta t,x)r(k,0) + b_2'(k\Delta t,x)r(k,x) + \sigma(k,k+x)\bar{\sigma}(k,k+x)\\
R(r(k,x),r(k,0)) &= b_0'(k\Delta t,0) + b_1'(k\Delta t,0)r(k,0) + b_2'(k\Delta t,0)r(k,x). \qquad (24)
\end{aligned}
$$

Equation (23) can be synthesized as

$$S_{k+1} - S_k = F_k(S_k;\theta)\Delta t + V_k(S_k;\theta)\Delta W_k, \tag{25}$$

with $F_k(\cdot)$ and $V_k(\cdot)$ corresponding to (21) and (22) respectively and where $\Delta W_k \sim \mathcal{N}(0,\Delta t)$. We have used the shorthand notation $r(k,x)$ to represent $r(k\Delta,x)$ and, similarly, with other quantities such as S_k standing for $S(k\Delta t)$.

The quantities we observe in the market are the LIBOR rates $l_x(t)$. From equations (9) and (15), we have that $l_x(t)$ is related to the state variables by

$$\ln l_x(t) = -\ln x + \bar{b}_0(t,x) + \bar{b}_1(t,x)r(t) + \bar{b}_2(t,x)r(t,x). \tag{26}$$

Thus it is most convenient to treat $\ln l_x(t)$ as the observations equation in our system.

Using Y_k to denote $\ln l_x(k\Delta t)$, the observation equation (26) becomes

$$Y_k = C_k S_k + \bar{b}_0(k\Delta t,x) - \ln x + q_k\eta_k, \tag{27}$$

where

$$C_k = (\bar{b}_1(k\Delta t,x), \bar{b}_2(k\Delta t,x)). \tag{28}$$

We have assumed in (27) the existence of an observation noise term $q_k\eta_k$, where $\eta_k \sim \mathcal{N}(0,1)$ is serially uncorrelated and independent of the ΔW_k. The strength of the observation noise, q_k, would reflect features (such as bid-ask spread) of the LIBOR market. In order to express the observation equation (27) in standard form we define the noise term

$$\lambda_k = q_k\eta_k, \tag{29}$$

so that

$$\lambda_k \sim N(0, \Lambda_k), \tag{30}$$

where

$$\Lambda_k = q_k^2. \tag{31}$$

With this notation the observation equation (27) may be written

$$Y_k = C_k S_k + \bar{b}_0(k\Delta t, x) - \ln x + \lambda_k. \tag{32}$$

4 The Dynamic Bayesian Updating Algorithm

From (25) we obtain for the conditional distribution of S_{k+1}, given S_k, the Gaussian distribution

$$p_\theta(S_{k+1}|S_k) \sim \mathcal{N}(S_{k+1}; S_k + F_k(S_k, \theta t)\Delta t, V_k(S_k, \theta t)V_k'(S_k, 0)\Delta t) \tag{33}$$

where we use the notation $\mathcal{N}(X; m, \Sigma)$ to denote a Gaussian random variable X with mean m and covariance matrix Σ. More specifically, we have

$$V_k(S_k, \theta)V_k'(S_k, \theta)\Delta t = \begin{pmatrix} g^2(r(k,0)\Delta t) & e^{-\lambda x}g^2(r(k,0)\Delta t) \\ e^{-\lambda x}g^2(r(k,0)\Delta t) & e^{-2\lambda x}g^2(r(k,0))\Delta t \end{pmatrix}, \tag{34}$$

which is immediately seen to be singular (the conditional correlation among the two components is equal to 1 and their joint distribution degenerates). The two components of the state vector are in fact linearly dependent and one can write

$$\begin{aligned} r(k+1,x) &= e^{-\lambda x}r(k+1,0) + [(r(k,x) + R_x(r(k,x), r(k,0)) - e^{-\lambda x}(r(k,0) + R(r(k,x), r(k,0)))] \\ &= \alpha r(k+1,0) + \beta(r(k,x), r(k,0)) \end{aligned} \tag{35}$$

thereby implicitly defining the constant α and the function $\beta(\cdot)$ that (see (24) with (11), (12) and (2)) is uniformly continuous in its arguments.

Putting $S_k = (S_k^1, S_k^2)$ where $S_k^1 = r(k,0)$, $S_k^2 = r(k,x)$, and making use of (35) we may also rewrite (25) as

$$\begin{aligned} S_{k+1}^1 &= S_k^1 + F_k^1(S_k, \theta)\Delta t + V_k^1(S_k, \theta)\Delta W_k, \\ S_{k+1}^2 &= \alpha S_{k+1}^1 + \beta(S_k), \end{aligned} \tag{36}$$

where F_k^1 and V_k^1 are the first components of the 2-vectors F_k and V_k in (25) respectively. Below we shall also consider the conditional distribution of S_{k+1}^1, given S_k, that is induced by (36) namely

$$p_\theta(S_{k+1}^1|S_k) \sim \mathcal{N}(S_{k+1}^1; S_k^1 + F_k^1(S_k, \theta)\Delta t, (V_k^1(S_k, \theta))^2\Delta t) \tag{37}$$

and that admits a density with respect to the Lebesgue measure on $\mathbb{R}^1$ (notice that $p_\theta(S_{k+1}|S_k)$, being degenerate, does not admit a density with respect to the Lebesgue measure on $\mathbb{R}^2$).

On the other hand, from (32) we obtain the conditional distribution of Y_k given S_k and θ as

$$p(Y_k|S_k, \theta) = \frac{1}{\sqrt{2\pi\Lambda_k}} \exp\left\{ -\frac{1}{2\Lambda_k} \left[Y_k - C_k S_k - \bar{b}_0(k\Delta t, x) + \ln x \right]^2 \right\}. \tag{38}$$

Using the representation (36) of the dynamics (25) we may also consider the distribution of Y_k given S_k^1, S_{k-1} and θ namely

$$\bar{p}\left(Y_k|S_k^1, S_{k-1};\theta\right) = \frac{1}{\sqrt{2\pi\Lambda_k}}\exp\left\{-\frac{1}{2\Lambda_k}\left[Y_k - \bar{b}_1(k\Delta t, x)S_k^1 - \bar{b}_2(k\Delta t, x)\alpha S_k^1\right.\right.$$
$$\left.\left. -\bar{b}_2(k\Delta t, x)\beta(S_{k-1}) - \bar{b}_0(k\Delta t, x) + \ln x\right]^2\right\}. \tag{39}$$

We are interested in the conditional joint distribution $p(S_k, \theta|y^k)$ of S_k and θ in the generic period $k = k\Delta t$, given the observations $y^k = (y_1, ..., y_k)$. Notice that, since the two components of S_k are linearly dependent, $p(S_k, \theta|y^n)$ does not in general have a density with respect to the Lebesgue measure on $\mathbb{R}^2$. However, it still satisfies the recursive Bayes formula, namely

$$p(S_{k+1}, \theta|y^{k+1}) \propto p(Y_{k+1}|S_{k+1}, \theta)\int p_\theta(S_{k+1}|S_k)dp(S_k, \theta|y^k) \tag{40}$$

with initial condition that we choose of the form $p(S_0, \theta) = p(S_0)p(\theta)$ (independence of S_0 and θ) and where $\propto$ denotes "proportional to".

In the Bayesian estimation procedure the parameter vector θ is assumed to take only a finite number of values in a hypercube whose upper and lower bounds would be specified by economic considerations of the range of likely parameter values. So in our context this vector is considered as a discrete random variable.

Since S_k takes a continuum of possible values (its dynamics are driven by the Gaussian ΔW_k), to actually compute the recursion (40) we discretize the values of S_k (for the convengence of the ensuing approximation see Proposition 4.1 below).

For this purpose, given a step-size $\delta > 0$ and an integer H, consider the square in $\mathbb{R}^2$ given by $(-H\delta, H\delta] \times (-H\delta, H\delta]$ and its partition (grid) into $4H^2$ squares, the generic l-th $(1 \leq l \leq 4H^2)$ of which is derived as follows: let

$$l = (2H)h + k + 1$$

where $0 \leq h < 2H$, $0 \leq k < 2H$, then

$$R^l = ((-H+k)\delta, (-H+k+1)\delta) \times ((-H+h)\delta, (-H+h+1)\delta))$$
$$:= (\alpha_1^l, \beta_1^l] \times (\alpha_2^l, \beta_2^l] \tag{41}$$

thereby implicitly defining $\alpha_1^l, \beta_1^l, \alpha_2^l, \beta_2^l$. In each of the $4H^2$ squares R^l pick a representative element, e.g. its middle point, namely

$$\left(\left(-H+k+\frac{1}{2}\right)\delta, \left(-H+h+\frac{1}{2}\right)\delta,\right) \tag{42}$$

if $l = (2H)h + k + 1$ (alternative choices for representative elements are equally valid).

In addition to the squares R^l for $l \leq 4H^2$ that form a partition of $(-H\delta, H\delta] \times (-H\delta, H\delta]$, consider further 8 subsets of $\mathbb{R}^2$, denoted also by R^l, and that for $l = 4H^2+1, ..., 4H^2+8$ are given by (on the right of each set there appears a possible choice for its representative element)

$$\begin{aligned}
R^{4H^2+1} &= (-\infty, -H\delta] \times (-\infty, -H\delta], && \left((-H-\frac{1}{2})\delta, (-H-\frac{1}{2})\delta\right) \\
R^{4H^2+2} &= (-H\delta, H\delta] \times (-\infty, -H\delta], && \left(0, (-H-\frac{1}{2})\delta\right) \\
R^{4H^2+3} &= (H\delta, +\infty] \times (-\infty, -H\delta], && \left((H+\frac{1}{2})\delta, (-H-\frac{1}{2}\delta)\right) \\
R^{4H^2+4} &= (-\infty, -H\delta] \times (-H\delta, H\delta], && \left((-H-\frac{1}{2})\delta, 0\right) \\
R^{4H^2+5} &= (H\delta, +\infty] \times (-H\delta, H\delta], && \left((H+\frac{1}{2})\delta, 0\right) \\
R^{4H^2+6} &= (-\infty, -H\delta] \times (H\delta, +\infty], && \left((-H-\frac{1}{2})\delta, (H+\frac{1}{2})\delta\right) \\
R^{4H^2+7} &= (-H\delta, H\delta] \times (H\delta, +\infty], && \left(0, (H+\frac{1}{2})\delta\right) \\
R^{4H^2+8} &= (H\delta, \infty] \times (H\delta, +\infty]. && \left((H+\frac{1}{2})\delta, (H+\frac{1}{2})\delta\right)
\end{aligned}$$

We shall use the notation $R^l = (\alpha_1^l, \beta_1^l] \times (\alpha_2^l, \beta_2^l)$ also for $l = 4H^2+1, ..., 4H^2+8$. Notice that $R^l(l = 1, ...4H^2+8)$ forms now a partition of all of $\mathbb{R}^2$ and we shall denote the chosen representative element of R^l by $\underline{r}^l = (r_1^l, r_2^l)$.

Given the discrete-time, continuous state Markov chain S_k with transition kernel $p_\theta(S_{k+1}|S_k)$, consider next the discrete-time finite state Markov chain $\bar{S}_k$, induced by S_k and having state space $\underline{r}^l(l = 1, ..., 4H^2+8)$. Denote by $P_{[i,h]}^{(k,\theta)}$ the generic (i, h)-th element $(i, h = 1, ...4H^2+8)$ of the corresponding transition probability matrix in period k. We then have, using the explicit representation (23) of the state transition equation (25),

$$
\begin{aligned}
P^{(k,\theta)}_{[i,h]} &= p_\theta\left\{\bar{S}_{k+1} = \underline{r}^h | \bar{S}_k = \underline{r}^i\right\} = P_\theta\left\{\left(S^1_{k+1}, S^2_{k+1}\right) \in R^h | (r^i_1, r^i_2)\right\} \\
&= P_\theta\{(r(k+1,x), r(k+1,0)) \in R^h | (r(k,x), r(k,0)) = (r^i_1, r^i_2)\} \\
&= P_\theta\left\{\alpha^h_1 \le r(k+1,x) < \beta^h_1, \alpha^h_2 \le r(k+1,0) < \beta^h_2 | r(k,x) = r^i_1, r(k,0) = r^i_2\right\} \\
&= P_\theta\left\{\frac{\alpha^h_1 - r^i_1 - R_x(r^i_1, r^i_2)\Delta}{e^{-\lambda x} g(r^i_2)} \le \Delta W_k \le \frac{\beta^h_1 - r^i_1 - R_x(r^i_1, r^i_2)\Delta}{e^{-\lambda x} g(r^i_2)}\right. , \\
&\qquad \left.\frac{\alpha^h_2 - r^i_2 - R(r^i_1, r^i_2)\Delta}{g(r^i_2)} \le \Delta W_k \le \frac{\beta^h_2 - r^i_2 - R(r^i_1, r^i_2)\Delta}{g(r^i_2)}\right\} \\
&= P_\theta\left\{\max\left[\frac{\alpha^h_1 - r^i_1 - R_x(r^i_1, r^i_2)\Delta}{e^{-\lambda x} g(r^i_2)}, \frac{\alpha^h_2 - r^i_2 - R(r^i_1, r^i_2)\Delta}{g(r^i_2)}\right] \le \Delta W_k\right. \\
&\qquad \left.\le \min\left[\frac{\beta^h_1 - r^i_1 - R_x(r^i_1, r^i_2)\Delta}{e^{-\lambda x} g(r^i_2)}, \frac{\beta^h_2 - r^i_2 - R(r^i_1, r^i_2)\Delta}{g(r^i_2)}\right]\right\} \\
&= \Phi\left(\frac{1}{\sqrt{\Delta}} \min\left[\frac{\beta^h_1 - r^i_1 - R_x(r^i_1, r^i_2)\Delta}{e^{-\lambda x} g(r^i_2)}, \frac{\beta^h_2 - r^i_2 - R(r^i_1, r^i_2)\Delta}{g(r^i_2)}\right]\right) \\
&\qquad - \Phi\left(\frac{1}{\sqrt{\Delta}} \max\left[\frac{\alpha^h_1 - r^i_1 - R_x(r^i_1, r^i_2)\Delta}{e^{-\lambda x} g(r^i_2)}, \frac{\alpha^h_2 - r^i_2 - R(r^i_1, r^i_2)\Delta}{g(r^i_2)}\right]\right)
\end{aligned} \tag{43}
$$

where $\Phi(\cdot)$ is the cumulative standard Gaussian distribution function.

The recursive Bayes' formula that corresponds to (40) for the discretized chain $\bar{S}_k$ becomes then

$$
\bar{p}\left(\bar{S}_{k+1} = \underline{r}^h, \theta | y^{k+1}\right) \propto p\left(Y_{k+1} | \bar{S}_{k+1} = \underline{r}^h, \theta\right) \sum_{i=1}^{4H^2+8} P^{(k,\theta)}_{[i,h]} \bar{p}(\bar{S}_k = \underline{r}^i, \theta | y^n), \tag{44}
$$

with initial condition $\bar{p}(S_0 = \underline{r}^i, \theta) = P\{S_0 \in R^i\} p_0(\theta)$ and the normalizing factor is the inverse of

$$
\sum_\theta \sum_{h=1}^{4H^2+8} p(Y_{k+1} | \bar{S}_{k+1} = \underline{r}^h, \theta) \cdot \sum_{i=1}^{4H^2+8} P^{(k,\theta)}_{[i,h]} \cdot \bar{p}(\bar{S}_k = \underline{r}^i, \theta | y^k). \tag{45}
$$

Formula (44) can be written in matrix form as

$$
\bar{p}^\theta_{k+1} \propto L^\theta_{k+1} (P^{(k,\theta)})^T \bar{p}^\theta_k, \tag{46}
$$

where $\bar{p}^\theta_k$ is the $(4H^2+8)$-vector with entries

$$
\bar{p}(\bar{S}_k = \underline{r}^i; \theta | y^k), \quad (i = 1, \cdots, 4H^2+8),
$$

$P^{(k,\theta)}$ is the matrix with elements $P^{(k,\theta)}_{[i,h]}$ $(i, h = 1, \cdots, 4H^2+8)$ and L^θ_{k+1} is the $(4H^2+8) \times (4H^2+8)$-diagonal matrix with entry in position h $(h = 1, \cdots, 4H^2+8)$ given by (see (38))

$$
\begin{aligned}
&p(Y_{k+1} | \bar{S}_{k+1} = \underline{r}^h, \theta) = \frac{1}{\sqrt{2\Lambda_{k+1}}} \cdot \\
&\exp\{-\frac{1}{2\Lambda_{k+1}}(Y_{k+1} - \bar{b}_2((k+1)\Delta t, x) r^h_1 - \bar{b}_1((k+1)\Delta t, x) r^h_2 - \bar{b}_0((k+1)\Delta t, x) + \ln x)^2\}.
\end{aligned} \tag{47}
$$

From the joint conditional distribution $\bar{p}(\bar{S}_k = \underline{r}^h, \theta | y^k)$, $(h = 1, \cdots, 4H^2 + 8)$ one can obtain the marginal conditional distributions

$$\bar{p}(\bar{S}_k = \underline{r}^h | y^k) = \sum_{\theta} \bar{p}(\bar{S}_k = \underline{r}^h, \theta | y^k), \tag{48}$$

$$\bar{p}(\theta | y^k) = \sum_{h=1}^{4H^2+8} \bar{p}(\bar{S}_k = \underline{r}^h, \theta | y^k), \tag{49}$$

of $\bar{S}_k$ and θ respectively.

Combining (48) with (44) and (45) we obtain the explicit expression

$$\bar{p}(\bar{S}_k = \underline{r}^h | y^k) = \frac{\sum_{\theta} p(Y_{k+1} | \bar{S}_{k+1} = \underline{r}^h, \theta) \sum_{i=1}^{4H^2+8} P_{[i,h]}^{(k,\theta)} \cdot \bar{p}(\bar{S}_k = i, \theta | y^k)}{\sum_{\theta} \sum_{h=1}^{4H^2+8} p(Y_{k+1} | \bar{S}_{k+1} = \underline{r}^h, \theta) \sum_{i=1}^{4H^2+8} P_{[i,h]}^{(k,\theta)} \bar{p}(\bar{S}_k = i, \theta | y^k)} \tag{50}$$

and, analogously, for $\bar{p}(\theta | y^k)$.

We next show that the discretization introduced above to make the recursion (40) computable is meaningful by showing that the approximate conditional distributions computed via (44) converge in a suitable weak sense to the original conditional distribution corresponding to (40). Since θ is discrete already from the outset, it suffices that we consider the convergence of the conditional distributions for each fixed value of θ. We have in fact the following

Proposition 4.1 *(Weak convergence of conditional distributions)*

Given any uniformly continuous and bounded function $F(S)$ on $\mathbb{R}^2$, we have for any period k, for any sequence of observations y^k and for any of the finite values of θ

$$\begin{aligned} \lim_{\substack{H \to \infty \\ \delta \to 0}} & \sum_{h=1}^{4H^2+8} F(\underline{r}^h) \bar{p}(\underline{r}^h, \theta | y^k) \\ &= \int F(S_k) dp(S_k, \theta | y^k) dS_k, \end{aligned} \tag{51}$$

where $\bar{p}(\underline{r}^h, \theta | y^k)$ is as in (44) and $p(S_k, \theta | y^k)$ as in (40).

Proof: See Appendix 2.

From Proposition 4.1, we immediately obtain the following corollary (for (i) below take $F \equiv 1$).

Corollary 4.1 *(Convergence of the marginal distributions of θ and weak convergence of the* *marginal distributions of S_k).*

$$(i) \quad \lim_{\substack{H \to \infty \\ \delta \to 0}} \bar{p}(\theta|y^k) = \lim_{\substack{H \to \infty \\ \delta \to 0}} \sum_{h=1}^{4H^2+8} \bar{p}(\underline{r}^h, \theta|y^k)$$

$$= \int dp(S_k, \theta|y^k) = p(\theta|y^k);$$

$$(ii) \quad \lim_{\substack{H \to \infty \\ \delta \to 0}} \sum_{h=1}^{4H^2+8} F(\underline{r}^h)\bar{p}(\underline{r}^h|y^k) = \lim_{\substack{H \to \infty \\ \delta \to 0}} \sum_{\theta} \sum_{h=1}^{4H^2+8} F(\underline{r}^h)\bar{p}(\underline{r}^h, \theta|y^k)$$

$$= \sum_{\theta} \int F(S_k) dp(S_k, \theta|y^k) = \int F(S_k) dp(S_k|y^k).$$

Remark:

Since power functions are not uniformly continuous nor bounded, Proposition 4.1 and Corollary 4.1 would not allow us to obtain convergence of the conditional moments. We can, however, obtain their convergence by truncating them with an arbitrarily large truncation factor.

5 Empirical Analysis

In this section we present numerical results for the Bayesian updating estimation procedure applied to real financial data. Figure 3 displays the monthly US LIBOR rates[4] for one month to twelve months quoted in the period from the 1st of December 1997 to the 13th of November 1998 with time (in days) on one axis and maturity (as fraction of a year) on the other axis. We chose this period as it seemed to be typical of periods not characterised by large interest rate volatility that might require more sophisticated treatment of the noise term, such as stochastic volatility or jump processes.

We observe from Figure 3 that the short maturity LIBOR rates (say less than 3-months) seem to undergo very little diffusion type movement. This is probably due to the fact that the short maturities of the yield curve are strongly influenced by the interest rate policy of the Federal Reserve. So it is unlikely that their motion is fully captured by the stochastic differential equation systems discussed in Section 2. To properly capture the dynamics of the short rate one would really need to also incorporate the interest rate policy function of the Federal Reserve and this is beyond the scope of the current paper. For this reason we perform our estimation procedure using only the $x = 0.5$ and $x = 1$ time series for the LIBOR rates that we display in Figure 4(a) and for which we have 250 observations. We have in fact, found that the results obtained for the $x = 0.25$ time series are not consistent with those obtained for the other values of x.

The observed initial forward curve, Figure 4b, is formed using the 1 to 12 months LIBOR rates and the 1 to 15 year swap rates at the beginning of the period. This curve is approximated by

$$r(0, x) = a_0 + (b_0 + b_1 x + b_2 x^2)e^{-b_3 x}, \tag{52}$$

[4]Our data source was Data Stream$^{\text{TM}}$.

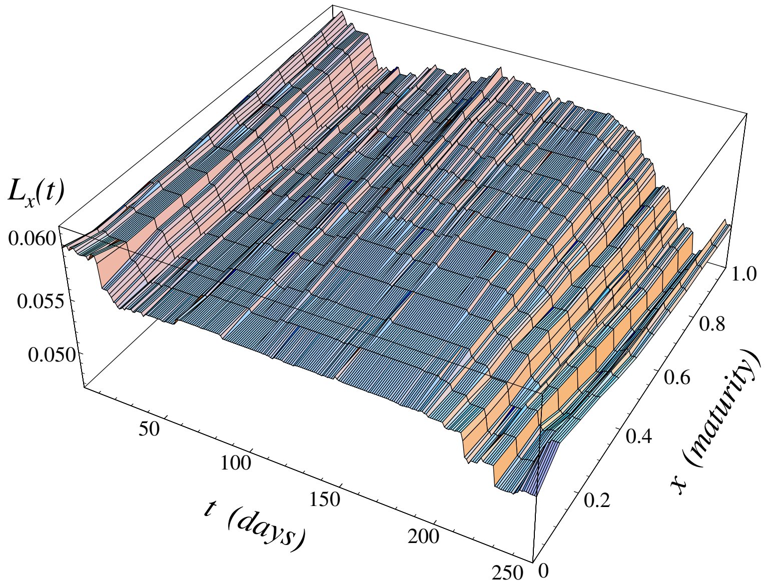

Figure 3: Monthly US LIBOR rates. Daily from December 1997 to November 1998

where $a_0 = 0.0677012, b_0 = -0.0087993, b_1 = 0.0001126, b_2 = -0.0001247$ and $b_3 = 0.1479538$ are obtained by least squares optimisation.

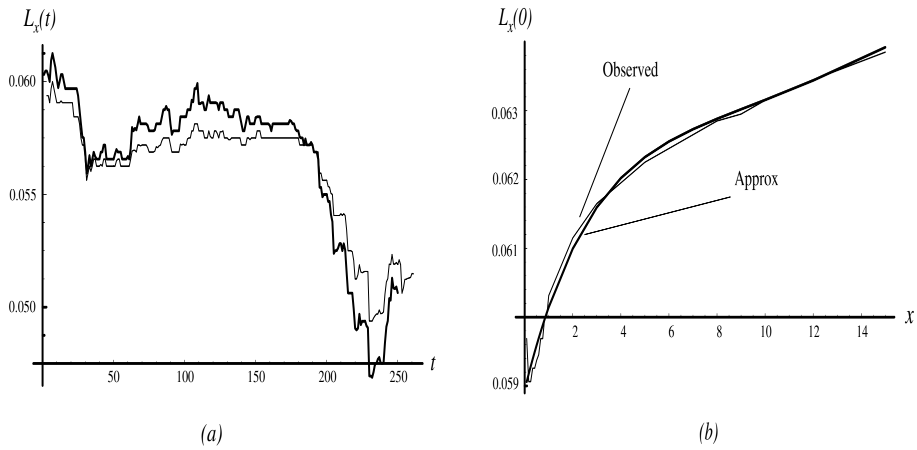

Figure 4: Showing $L_x(t)$, for $x = 0.5$ (thin line) and 1.0 (thick line) (a), and the initial forward curve (b).

The range of LIBOR rates observed over this period suggests that the state vector $R = (r(t,x), r(t,0))$ should lie in the square $[0.02, 0.09] \times [0.02, 0.09]$, which is discretised into 15×15 cells, each represented by its centre value. Also, surrounding this region are 8 cells, representing the remaining possible values of R. The possible ranges of values for $\theta = (\lambda, \sigma_0, \delta)$ are not known *a priori*. Rough estimates could be found by comparing various moments of observed changes in LIBOR rates with the expressions calculated from the model. However we have found the most effective way to determine the ranges of θ is by successive application of the Bayesian algorithm over a successively decreasing range and finer grid, until the support is found. This approach avoids the need to search over large regions of the parameter space where the support is zero. Table 1 shows the final lower and upper bounds found by this procedure for each parameter and the number of divisions used to achieve the results reported.

x	ΔL_x		λ			σ_0			δ		
	mean	std.	low	up	# div.	low	up	# div.	low	up	# div.
0.5	3.0E-5	2.7E-4	-2.45	2.45	80	0.0	3.0	70	1.0	3.2	50
1.0	3.9E-5	3.4E-4	-1.55	1.45	80	0.0	3.0	70	1.0	3.2	50

Table 1: Table showing the mean and standard deviation of the change in $L_x(t)$ and the ranges for the parameters.

Figures 5, 6 and 7 show the estimated marginal distributions for λ, σ_0 and δ respectively for $x = 0.5$ and $x = 1$. The mean and standard deviations of each distribution are shown in Table 2. These values may presumably be regarded as the "best" estimates for the parameters. The estimates for σ_0 and δ display some consistency across the two values of x. However the two distributions $p(\lambda)$ in Figure 5, whilst showing the same general shape clearly have different means. This is probably due to the fact that this parameter is the one that is most sensitive to the different evolutions of the six-month and 1-year LIBOR rates evident in Figure 4a.

The main comment to make in relation to the distributions for $p(\lambda)$ is that the great bulk of probability mass and the means are in the $\lambda < 0$ region. This indicates that over the short maturity (up to one year) of the data set, volatility is increasing with maturity. This seems to contradict the comment in footnote 2, though we did add the qualification that volatility may increase with maturity over very short maturities. In fact, using an HJM model for the dynamics of futures prices and estimating it by a maximum likelihood method using futures price data, Bhar, Chiarella and Tô (2002) have found that there is a hump in the volatility function at about one year. This is consistent with our negative estimates of λ in the present study. The increasing volatility of the LIBOR rates with maturity can also be seen by casual inspection of Figure 3 where the time series of $L_x(t)$ at $x = 1.0$ seems to have more variation than that at $x = 0.5$.

Looking at the three marginal distributions as a whole, one starts to appreciate why point parameter estimates vary quite a deal amongst the various empirical studies cited in the introduction. First the estimate of λ is clearly affected by the maturity of the data series used. The estimates of σ_0 and δ are far less sensitive to maturity, but their distributions are quite wide,

x	λ		σ_0		δ	
	mean	std. dev.	mean	std. dev.	mean	std. dev.
0.5	-0.913	0.902	1.449	0.89	2.325	0.343
1.0	-0.297	0.311	1.459	0.884	2.189	0.316

Table 2: Mean and standard deviation from the distributions of the parameters.

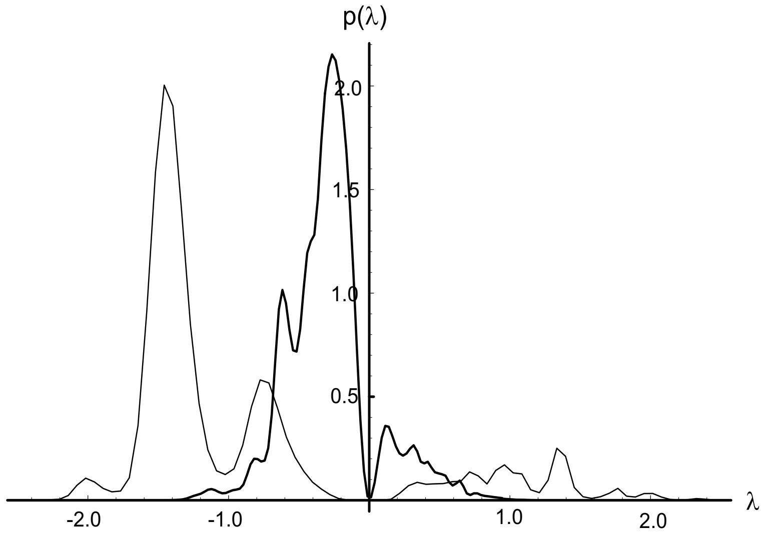

Figure 5: The marginal distribution $p(\lambda)$ for $x = 0.5$ (thin line) and $x = 1$ (thick line).

implying that a wide range of point estimates are equally likely [5]. Certainly Figure 7 is consistent with the wide range of point estimates of δ obtained in the empirical literature, however the marginal distribution $p(\delta)$ has its mean at a much higher value (> 2) than the point estimates (0.5 to 1.5) reported in much empirical literature. We believe the higher values of δ may be a feature of LIBOR rates. Using a similar model structure for LIBOR rate dynamics to the one of this paper, but a quite different estimation approach and a longer LIBOR rate series, Chiarella, Hung and Tô (2005) obtained a point estimate of δ around 2, which is quite consistent with the mean values reported in Table 2.

6 Conclusion

We have derived the risk-neutral dynamics for unobserved factors upon which pure discount bond prices depend within the Heath-Jarrow-Morton framework using a certain forward rate volatility specification. We have then used the link between LIBOR rates, forward rates and pure discount bond prices to obtain the corresponding dynamics for LIBOR rates. The overall stochastic dynamic system can then be treated as a partially observed system with changes in the LIBOR rates being the observations. Since data are observed discretely, we have considered a discretised version of the model and developed a dynamic Bayesian updating algorithm to compute the posterior distribution for the model parameters conditional on the observed market LIBOR rates. The algorithm has been applied to some U.S. data and gives a model fit that seems consistent with some of the traditional econometric studies. The estimated marginal distributions of the parameters help to explain the wide range of point estimates for some parameters obtained using traditional econometric estimation procedures. Our results also suggest

[5] We conjecture the probability density $p(\sigma_0)$ approaches a uniform distribution as the discretisation in σ_0 becomes finer.

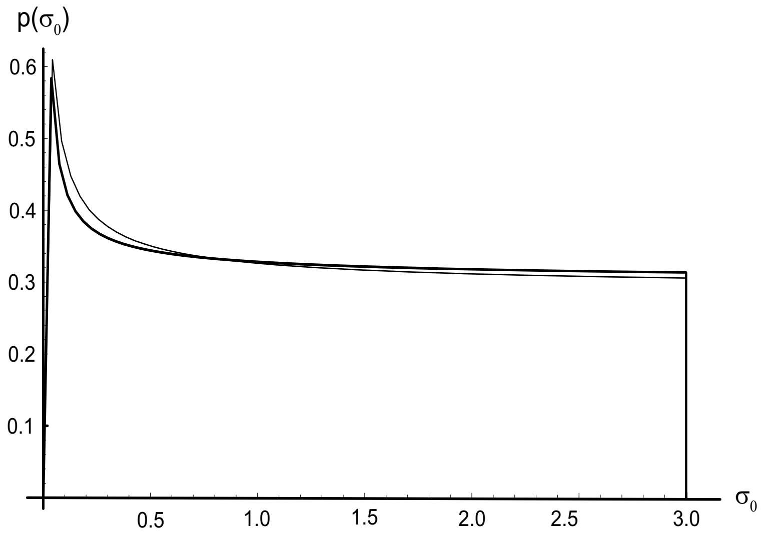

Figure 6: The marginal distribution $p(\sigma_0)$ for $x = 0.5$ (thin line) and $x = 1$ (thick line).

that the dynamics for LIBOR rates may be characterised by a higher interest rate elasticity (the parameter δ) than for the one–month Treasury rates used in much of the empirical literature.

Future research could go in a number of directions. First, there is a need to relax some of our restrictive assumptions, in particular allowing for more Wiener processes to drive the forward rate dynamics, since a number of empirical studies suggest that at least two and maybe three factors need to be considered. In this regard see Chiarella, Hung and Tô (2005). Second, one could exploit the fact that in addition to avoiding the use of proxy variables for the instantaneous spot rate of interest, the methodology proposed here has the advantage that a number of available discretely compounded rates may be used as the observed quantities. Thus it would be of interest to consider a data set with a range of maturities, particularly those corresponding to the most actively traded points on the yield curve. In this way one could obtain the volatility for the instantaneous spot rate most consistent with a set of discretely compounded LIBOR rates whose maturities are of most relevance to the application at hand. Third, more work also needs to be done on statistical diagnostics to assess the goodness-of-fit of the estimated models. Finally, more efficient numerical schemes for the implementation of the dynamic Bayesian updating need to be developed. Here we have relied on the Euler-Marayama discretisation of the stochastic dynamics and the attendant convenience of working with normal distributions. Further developments could involve using higher order discretisation schemes or using an entirely different philosophy for the dynamic Bayesian updating, such as particle filters, see for example Bølviken and Storvik (2001) and Chib et al. (2002).

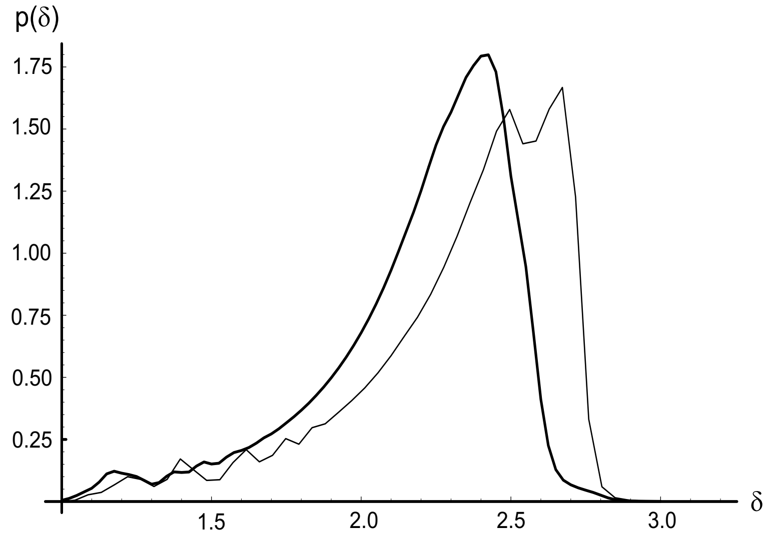

Figure 7: The marginal distribution $p(\delta)$.

7 Appendix 1: BM Dynamics - Forward Rate Dependent Volatility Function

Consider the HJM model within the BM parameterisation. Under the risk neutral measure $\widetilde{\mathbb{P}}$ the forward rate $r(t,x)$ satisfies the stochastic integal equation

$$r(t,x) = r(0,t+x) + \int_0^t \sigma(s,t+x)\bar{\sigma}(s,t+x)ds + \int_0^t \sigma(s,t+x)d\widetilde{w}(s), \tag{53}$$

where

$$\bar{\sigma}(s,t+x) = \int_s^{t+x} \sigma(s,u)du. \tag{54}$$

From(53) with $x = 0$ we obtain the stochastic integal equation for the instantaneous spot rate $r(t)(= r(t,0))$

$$r(t) = r(0,t) + \int_0^t \sigma(v,t)\bar{\sigma}(v,t)dv + \int_0^t \sigma(v,t)d\widetilde{w}(v). \tag{55}$$

Here we consider volatility functions of the form

$$\sigma(t,u) = g\left(r(t,x_1),...,r(t,x_n)\right)e^{-\lambda(u-t)}. \tag{56}$$

The dynamics for each state variable $r(t,x_i)(i = 1,...,n)$ is obtained by setting $x = x_i$ in (53). We note that with the specification (56), the expression (54) for $\bar{\sigma}(s,t)$ becomes

$$\bar{\sigma}(s,t) = \frac{(1-e^{-\lambda(t-s)})}{\lambda} g\left(r(s,x_1),...,r(s,x_n)\right). \tag{57}$$

To ease the notation we set

$$g(\vec{r}(s,\cdot)) \equiv g(r(s,x_1),...,r(s,x_n)), \tag{58}$$

so that we can write

$$\sigma(t,u) = g(\vec{r}(t,\cdot))\, e^{-\lambda(u-t)},$$

and

$$\bar{\sigma}(t,u) = g(\vec{r}(t,\cdot))\, \frac{1-e^{\lambda(u-t)}}{\lambda}.$$

With these various notations the stochastic integal equation for $r(t,x_i)$ becomes

$$\begin{aligned} r(t,x_i) = r(0,t+x_i) &+ \int_0^t g^2(\vec{r},(s,\cdot))\, e^{-\lambda(t+x_i-s)} \frac{1-e^{-\lambda(t+x_i-s)}}{\lambda} ds \\ &+ \int_0^t g(\vec{r}(s,\cdot))\, e^{-\lambda(t+x_i-s)} d\widetilde{w}(s), \end{aligned} \tag{59}$$

for $i = 1.2, ..., n.$

Note that in terms of the Chiarella and Kwon (2003) notation we may write

$$\sigma(t,u) = c_{11}(t)\sigma_{11}(u), \tag{60}$$

where

$$\sigma_{11}(u) = e^{-\lambda u}, \quad c_{11}(t) = e^{\lambda t} g(\vec{r}(t,\cdot)), \tag{61}$$

and so

$$\bar{\sigma}_{11}(x) = \frac{1-e^{-\lambda x}}{\lambda}. \tag{62}$$

The naturally arising subsidiary variables needed to Markovianise the dynamics then turn out to be[6]

$$\psi(t) = \int_o^t e^{\lambda s} g(\vec{r}(s,\cdot))\, d\widetilde{w}(s) - \int_o^t g(\vec{r}(s,\cdot))^2\, e^{2\lambda s} \frac{\left(1-e^{-\lambda s}\right)}{\lambda} ds, \tag{63}$$

$$\varphi(t) = \int_o^t c_{11}(s)^2 ds = \int_o^t g(\vec{r}(s,\cdot))^2\, e^{2\lambda s} ds. \tag{64}$$

The quantities $\psi(t)$, $\varphi(t)$ are both stochastic, so in terms of Chiarella and Kwon (2003) notation the stochastic differential equation for $r(t,x)$ is written

$$r(t,x) = r(0,t+x) + \sigma_{11}(t+x)\psi(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\varphi(t). \tag{65}$$

[6]Note that in Chiarella and Kwon (2003) $\psi(t)$, $\varphi(t)$ are respectively denoted $\psi_1'(t)$ and $\varphi_{11}(t)$.

By choosing any two values for x e.g. x_1, x_2, we have two equations for $\psi(t)$ and $\varphi(t)$ in terms of $r(t,x_1)$ and $r(t,x_2)$.

Thus we have the system

$$\begin{bmatrix} \sigma_{11}(t+x_1) & \sigma_{11}(t+x_1)\bar{\sigma}_{11}(t+x_1) \\ \sigma_{11}(t+x_2) & \sigma_{11}(t+x_2)\bar{\sigma}_{11}(t+x_2) \end{bmatrix} \begin{bmatrix} \psi(t) \\ \varphi(t) \end{bmatrix} = \begin{bmatrix} r(t,x_1) - r(0,t+x_1) \\ r(t,x_2) - r(0,t+x_2) \end{bmatrix},$$

so that

$$\begin{bmatrix} \psi(t) \\ \varphi(t) \end{bmatrix} = \begin{bmatrix} \Delta_{11}(t) & \Delta_{12}(t) \\ \Delta_{21}(t) & \Delta_{22}(t) \end{bmatrix} \begin{bmatrix} r(t,x_1) - r(0,t+x_1) \\ r(t,x_2) - r(0,t+x_2) \end{bmatrix}, \tag{66}$$

where

$$\boxed{\begin{aligned} &\Delta_{11}(t) = \sigma_{11}(t+x_2)\bar{\sigma}_{11}(t+x_2)/\Delta, \quad \Delta_{12}(t) = -\sigma_{11}(t+x_1)\bar{\sigma}_{11}(t+x_1)/\Delta, \\ &\Delta_{21}(t) = -\sigma_{11}(t+x_2)/\Delta, \quad \Delta_{22}(t) = \sigma_{11}(t+x_1)/\Delta, \\ &\Delta = \sigma_{11}(t+x_1)\sigma_{11}(t+x_2)\left(\bar{\sigma}_{11}(t+x_2) - \bar{\sigma}_{11}(t+x_1)\right). \end{aligned} \quad (67)}$$

Rearranging (66) we can express $\psi(t)$ and $\varphi(t)$ as

$$\psi(t) = \Delta_{11}(t)r(t,x_1) + \Delta_{12}(t)r(t,x_2) - \delta_1(t), \tag{68}$$
$$\varphi(t) = \Delta_{21}(t)r(t,x_1) + \Delta_{22}(t)r(t,x_2) - \delta_2(t), \tag{69}$$

where

$$\begin{aligned} \delta_1(t) &= \Delta_{11}(t)r(0,t+x_1) + \Delta_{12}(t)r(0,t+x_2), \\ \delta_2(t) &= \Delta_{21}(t)r(0,t+x_1) + \Delta_{22}(t)r(0,t+x_2). \end{aligned}$$

Thus the forward rate of any tenor can be expressed in terms of $r(t,x_1)$ and $r(t,x_2)$ by substituting (68) and (69) into (65) i.e.

$$\begin{aligned} r(t,x) = {} & r(0,t+x) - \sigma_{11}(t+x)\delta_1(t) - \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\delta_2(t) \\ & + \left[\sigma_{11}(t+x)\Delta_{11}(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\Delta_{21}(t)\right] r(t,x_1) \\ & + \left[\sigma_{11}(t+x)\Delta_{12}(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\Delta_{22}(t)\right] r(t,x_2). \end{aligned} \tag{70}$$

In terms of equation (22) of Chiarella and Kwon (2003) we have

$$\boxed{\begin{aligned} b_0(t,x) &= r(0,t+x) - \sigma_{11}(t+x)\delta_1(t) - \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\delta_2(t), \\ b_1(t,x) &= \sigma_{11}(t+x)\Delta_{11}(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\Delta_{21}(t), \\ b_2(t,x) &= \sigma_{11}(t+x)\Delta_{12}(t) + \sigma_{11}(t+x)\bar{\sigma}_{11}(t+x)\Delta_{22}(t). \end{aligned} \quad (71)}$$

Thus (70) can be written

$$r(t,x) = b_0(t,x) + \sum_{i=1}^{2} b_i(t,x)r(t,x_i). \tag{72}$$

We can then use equation (25) from Chiarella and Kwon (2003) to obtain the expression for the BM bond price, namely

$$\boxed{b(t,x) = exp(-[\bar{b}_o(t,x) + \sum_{i=1}^{2} \bar{b}_i(t,x) r(t,x_i)]), \qquad (73)}$$

where

$$\bar{b}_i(t,x) = \int_0^x b_i(t,u)du, \qquad (i = 0,1,2). \tag{74}$$

The $x-$ tenor LIBOR rate is related to $b(t,x)$ via

$$L_x(t) = \frac{1}{x}\left(\frac{1}{b(t,x)} - 1\right). \tag{75}$$

We treat (75) as the observation equation with the state variables $r(t,x_1),...,r(t,x_n)$ being driven the stochastic differential equation system (see Chiarella and Kwon (2003) just below equation (23)

$$\begin{aligned} dr(t,x_k) = &\left[b_0^{'}(t,x_k) + \sum_{i=1}^{2} b_i^{'}(t,x_k) r(t,x_i) + \sigma(t,t+x_k)\bar{\sigma}(t,t+x_k)\right] dt \\ &+ e^{-\lambda x_k} g(r(t,x_1),...,r(t,x_n)) d\widetilde{w}(t), \qquad (k = 1,2,...,n). \end{aligned} \tag{76}$$

Note that in (76)

$$b_i^{'}(t,x) \equiv \frac{\partial b_i(t,x)}{\partial x}. \tag{77}$$

From the particular specification in (60) we have

$$\sigma_{11}(t+x) = e^{-\lambda(t+x)}, \tag{78}$$

so that

$$\frac{\partial \sigma_{11}(t+x)}{\partial x} = -\lambda e^{-\lambda(t+x)} = -\lambda \sigma_{11}(t+x), \tag{79}$$

and

$$\bar{\sigma}_{11}(t+x) = \frac{1 - e^{-\lambda(t+x)}}{\lambda} = \frac{1 - \sigma_{11}(t+x)}{\lambda}, \tag{80}$$

$$\bar{\sigma}'_{11}(t+x) = -\frac{\sigma'_{11}(t+x)}{\lambda} = \sigma_{11}(t+x). \tag{81}$$

From equation (71) we calculate

$$\begin{aligned} b_0'(t,x) &= r'(0,t+x) - \sigma_{11}'(t+x)\left[\delta_1(t) + \bar{\sigma}_{11}(t+x)\delta_2(t)\right] - \bar{\sigma}_{11}'(t+x)\sigma_{11}(t+x)\delta_2(t) \\ &= r'(0,t+x) + \lambda\sigma_{11}(t+x)\left[\delta_1(t) + \bar{\sigma}_{11}(t+x)\delta_2(t)\right] - \sigma_{11}^2(t+x)\delta_2(t) \\ &= r'(0,t+x) + \sigma_{11}(t+x)\left[\lambda\delta_1(t) + (1-2\sigma_{11}(t+x))\delta_2(t)\right], \qquad (82) \\ b_1'(t,x) &= \sigma_{11}'(t+x)\Delta_{11}(t) + \sigma_{11}'(t+x)\bar{\sigma}_{11}(t+x)\Delta_{21}(t) + \bar{\sigma}_{11}'(t+x)\sigma_{11}(t+x)\Delta_{21}(t) \\ &= -\lambda\sigma_{11}(t+x)[\Delta_{11}(t) + \bar{\sigma}_{11}(t+x)\Delta_{21}(t)] + \sigma_{11}^2(t+x)\Delta_{21}(t) \\ &= \sigma_{11}(t+x)[-\lambda\Delta_{11}(t) - \lambda\bar{\sigma}_{11}(t+x)\Delta_{21}(t) + \sigma_{11}(t+x)\Delta_{21}(t)] \\ &= \sigma_{11}(t+x)[-\lambda\Delta_{11}(t) + (\sigma_{11}(t+x)-1)\Delta_{21}(t) + \sigma_{11}(t+x)\Delta_{21}(t)] \\ &= \sigma_{11}(t+x)[(2\sigma_{11}(t+x)-1)\Delta_{21}(t) - \lambda\Delta_{11}(t)], \qquad (83) \end{aligned}$$

and

$$b_2'(t,x) = \sigma_{11}(t+x)[(2\sigma_{11}(t+x)-1)\Delta_{22}(t) - \lambda\Delta_{12}(t)]. \qquad (84)$$

To operationalise equation (73) we calculate according to equation (74) the quantities [7]

$$\begin{aligned} \bar{b}_0(t,x) &= \int_0^x \left[r(0,t+u) - \sigma_{11}(t+u)\delta_1(t) - \sigma_{11}(t+u)\bar{\sigma}_{11}(t+u)\delta_2(t)\right] du \\ &= \bar{r}(0,t+x) - \delta_1(t)\int_0^x e^{-\lambda(t+u)}du - \delta_2(t)\int_0^x \frac{e^{-\lambda(t+u)} - e^{-2\lambda(t+u)}}{\lambda}du \\ &= \bar{r}(0,t+x) + \frac{\lambda\delta_1(t)+\delta_2(t)}{\lambda^2}(\sigma_{11}(t+x) - \sigma_{11}(t)) + \frac{\delta_2(t)}{2\lambda^2}(\sigma_{11}^2(t) - \sigma_{11}^2(t+x)), \\ & \qquad (85) \\ \bar{b}_1(t,x) &= \int_0^x \sigma_{11}(t+u)\Delta_{11}(t) + \sigma_{11}(t+u)\bar{\sigma}_{11}(t+u)\Delta_{21}(t)du \\ &= \int_0^x \sigma_{11}(t+u)\Delta_{11}(t) + \frac{\sigma_{11}(t+u) - \sigma_{11}^2(t+u)}{\lambda}\Delta_{21}(t)du \\ &= \frac{\lambda\Delta_{11}(t)+\Delta_{21}(t)}{\lambda^2}(\sigma_{11}(t) - \sigma_{11}(t+x)) + \frac{\Delta_{21}(t)}{2\lambda^2}(\sigma_{11}^2(t+x) - \sigma_{11}^2(t)), \qquad (86) \end{aligned}$$

and

$$\bar{b}_2(t,x) = \frac{\lambda\Delta_{12}(t)+\Delta_{22}(t)}{\lambda^2}(\sigma_{11}(t) - \sigma_{11}(t+x)) + \frac{\Delta_{22}(t)}{2\lambda^2}(\sigma_{11}^2(t+x) - \sigma_{11}^2(t)). \qquad (87)$$

For our application $x_1 = 0$ and $x_2 = x$ so in equation (67) we set $x_1 = 0, x_2 = x$, to obtain

$$\begin{aligned} \Delta_{11}(t) &= e^{2\lambda t}\frac{1-e^{-\lambda(t+x)}}{1-e^{\lambda x}}, \quad \Delta_{12}(t) = -e^{\lambda(2t+x)}\frac{1-e^{-\lambda t}}{1-e^{-\lambda x}}, \\ \Delta_{21}(t) &= -e^{2\lambda t}\frac{\lambda}{1-e^{-\lambda x}}, \quad \Delta_{22}(t) = e^{\lambda(2t+x)}\frac{\lambda}{1-e^{-\lambda x}}. \end{aligned}$$

From equations (12) and (54) we find that

$$\bar{\sigma}(t,t+x) = \frac{\sigma_0 r^\delta}{\lambda}(1-e^{-\lambda x}).$$

[7] Note that $\bar{r}(0,t+x) \equiv \int_0^x r(0,t+u)du$

8 Appendix 2: Proof of Proposition 4.1.

(The proof is an adaptation to the present situation of the one of Theorem 4.1 in Bhar, Chiarella and Runggaldier (2002).)

In the proof we shall use S_k^H to denote, for the generic period $k = k\Delta t$, the random variable previously denoted by $\bar{S}_k$ that takes the values $\underline{r}^h$ $(h = 1, \cdots, 4H^2 + 8)$ of the representative elements in the discretization defined in this section and has distribution $\bar{p}(\bar{S}_k, \theta | y^k)$ (see (44)) for a given value of θ. The proof then amounts to showing that, for any uniformly continuous function F,

$$\lim_{\substack{H \to \infty \\ \delta \to 0}} E_\theta\{F(S_k^H)|y^k\} = E_\theta\{F(S_k)|y^k\}. \tag{88}$$

We proceed by induction on k. For $k = 0$ we have by construction that

$$\lim_{\substack{H \to \infty \\ \delta \to 0}} E_\theta\{F(S_0^H)\} = E_\theta\{F(S_0)\}. \tag{89}$$

Assume then that the statement is true for k. In period $(k+1)\Delta t$ consider now (see (44))

$$\begin{aligned} &E_\theta\{F(S_{k+1}^H)|y^{k+1}\} = \sum_{h=1}^{4H^2+8} F(\underline{r}^h)\bar{p}(S_{k+1}^H = \underline{r}^h|y^{k+1}) \\ &\propto \sum_{h=1}^{4H^2+8} F(\underline{r}^h)p(y_{k+1}|\underline{r}^h, \theta) \sum_{i=1}^{4H^2+8} P_{[i,h]}^{(k,\theta)}\bar{p}(\underline{r}^i, \theta|y^k) \\ &= \sum_{i=1}^{4H^2+8} \phi_F^H(\underline{r}^i, \theta, y_{k+1})\bar{p}(\underline{r}^i, \theta|y^k) = E\{\phi_F^H(S_k^H, \theta, y_{k+1})|y^k\}. \end{aligned} \tag{90}$$

With

$$\begin{aligned} &\phi_F^H(\underline{r}^i, \theta, y_{k+1}) = \sum_{h=1}^{4H^2+8} F(\underline{r}^h)p(y_{k+1}|\underline{r}^h, \theta)P_{[i,h]}^{(k,\theta)} \\ &= \sum_{h=1}^{4H^2+8} F(\underline{r}^h)p(y_{k+1}|\underline{r}^h, \theta)\int_{R^h} dp_\theta(S_{k+1}|S_k^H = \underline{r}^i) \end{aligned} \tag{91}$$

where the rightmost equality follows from (43). Notice that

$$E_\theta\{F(S_{k+1}^H)|y^{k+1}\} = \frac{E_\theta\{\phi_F^H(S_k^H, \theta, y_{k+1}|y^k\}}{E_\theta\{\phi_1^H(S_k^H, \theta, y_{k+1})|y^k\}}. \tag{92}$$

On the other hand, according to (40) we have

$$\begin{aligned} &E_\theta\{F(S_{k+1})|y^{k+1}\} = \int F(S_{k+1})dp(S_{k+1}, \theta|y^{k+1}) \\ &\propto \int F(S_{k+1})p(y_{k+1}|S_{k+1}, \theta)\int dp_\theta(S_{k+1}|S_k)dp(S_k, \theta|y^k) \\ &= \int \phi_F(S_k, \theta, y_{k+1})dp(S_k, \theta|y^k). \end{aligned} \tag{93}$$

With

$$\phi_F(S_k, \theta, y_{k+1}) = \int F(S_{k+1}) p(y_{k+1}|S_{k+1}, \theta) dp_\theta(S_{k+1}|S_k) \tag{94}$$

and notice that, here too,

$$E_\theta\{F(S_{k+1})|y^{k+1}\} = \frac{E_\theta\{\phi_F(S_k, \theta, y_{k+1}|y^k\}}{E_\theta\{\phi_1(S_k, \theta, y_{k+1})|y^k\}}. \tag{95}$$

Given our assumptions, the function $F(S)p(y|S, \theta)$ with $p(y|S, \theta)$ as in (38) is uniformly continuous in S for all values of y and θ. Notice, in fact, that $p(y|S, \theta)$ is continuous in S and its limit for S going to infinity is, uniformly in (y, θ) equal to zero. Given $\eta > 0$, there is thus H_F^o (depending on F and the parameters in the model such as x and Λ_k) such that, for $H > H_F^o$ and $\delta < (H_F^o)^{-1}$, one has in any period $k\Delta t$ and for all values of S_k, θ, y_{k+1},

$$\begin{aligned} & |\phi_F^H(S_k, \theta, y_{k+1}) - \phi_F(S_k, \theta, y_{k+1})| \\ \leq & \sum_{h=1}^{4H^2+8} \int_{R^h} |F(\underline{r}^h) p(y_{k+1}|\underline{r}^h, \theta) - F(S_{k+1}) p(y_{k+1}|S_{k+1}, \theta)| dp_\theta(S_{k+1}|S_k) \\ \leq & \quad \eta \sum_{h=1}^{4H^2+8} \int dp_\theta(S_{k+1}|S_k) = \eta. \end{aligned} \tag{96}$$

To complete the induction step we have to show that $\phi_F(S; \theta, y)$ is, for all values of (θ, y), uniformly continuous in S. In fact, if this is the case, then, given $\eta > 0$, there exists $H_F^1(y^{k+1}, \theta)$ (depending on F and the parameters in the model) such that, for $H > H_F^1$ and $\delta < (H_F^1)^{-1}$,

$$|E_\theta\{\phi_F(S_k^H, \theta, y_{k+1})|y^k\} - E_\theta\{\phi_F(S_k, \theta, y_{k+1})|y^k\} \leq \eta. \tag{97}$$

From (96) and (97) it then follows that

$$\begin{aligned} & |E_\theta\{\phi_F^H(S_k^H; \theta, y_{k+1})|y^k\} - E_\theta\{\phi_F(S_k; \theta, y_{k+1})|y^k\}| \\ \leq & |E_\theta\{\phi_F^H(S_k^H; \theta, y_{k+1})|y^k\} - E_\theta\{\phi_F(S_k^H; \theta, y_{k+1})|y^k\}| \\ & + |E_\theta\{\phi_F(S_k^H; \theta, y_{k+1})|y^k\} - E_\theta\{\phi_F(S_k; \theta, y_{k+1})| \leq 2\eta \end{aligned} \tag{98}$$

for $H > H_F := \max\{H_F^o, H_F^1(y^{k+1}, \theta)\}$ and $\delta < (H_F)^{-1}$. Combining (98) with (92) and (95) allows to complete the induction step and thus the proof of the proposition.

It remains thus to show that $\phi_F(S; \theta, y)$ is uniformly continuous in S for each pair (θ, y).

To this effect notice that, by (36) as well as (38) and (39) (see also (94)) we have that

$$\begin{aligned} \phi_F(S_k, \theta, y^{k+1}) &= \int F(S_{k+1}^1, S_{k+1}^2) p(y_{k+1}|S_{k+1}^1, S_{k+1}^2, \theta) \cdot dp_\theta(S_{k+1}^1, S_{k+1}^2|S_k^1, S_k^2) \\ &= \int \bar{F}(S_{k+1}^1, S_k^1, S_k^2) \bar{p}(y_{k+1}|S_{k+1}^1, S_k^1, S_k^2, \theta) \cdot p_\theta(S_{k+1}^1|S_k^1, S_k^2) dS_{k+1}^1, \end{aligned} \tag{99}$$

where

$$\bar{F}(S_{k+1}^1, S_k^1, S_k^2) := F(S_{k+1}^1, \alpha S_{k+1}^1 + \beta(S_k^1, S_k^2)) \tag{100}$$

and $p_\theta(S^1_{k+1}|S^1_k,S^2_k)$ coincides with the Gaussian density $p_\theta(S^1_{k+1}|S_k)$ in (37).

Since (see comment after (35)) $\beta(S^1_k,S^2_k)$ is uniformly continuous we have that $\bar F(S^1_{k+1},S^1_k,S^2_k)$ is bounded and uniformly continuous in $S_k=(S^1_k,S^2_k)$.

We next have from (37) with (35) and (22) that

$$p_\theta(S^1_{k+1}|S_k)=\frac{1}{\sqrt{2\pi}g(r(k,0))}\cdot\exp\left\{-\frac{1}{2g^2(r(k,0))}(S^1_{k+1}-S^1_k-R(r(k,x),r(k,0)))^2\right\}$$
$$:=\phi_1(S_k)\cdot\phi_2(S_k;S^1_{k+1}) \tag{101}$$

where (see (11) and (24)) $\phi_1(S_k)$ and $\phi_2(S_k;S^1_{k+1})$ are bounded functions that are uniformly continuous in S_k. Concerning $\phi_2(S_k;S^1_{k+1})$ notice in fact that it is continuous and the limit for S_k going to infinity is, uniformly in S^1_{k+1}, equal to zero. Being a product of two bounded and uniformly continuous functions, $p_\theta(S^1_{k+1}|S_k)$ is thus bounded and uniformly continuous in S_k.

Coming to $\bar p(y_{k+1}|S^1_{k+1},S_k;\theta)$ defined in (39), it is immediately seen that, as a function of S_k, it is bounded uniformly in (y_{k+1},S_{k+1},θ). Since it is continuous in S_k and the limit for S_k going to infinity is zero uniformly in (y_{k+1},S_{k+1},θ), it is also a uniformly continuous function of S_k. Notice, furthermore, that the integral with respect to S^1_{k+1} is equal to the constant $\gamma=\bar{b_1}((k+1)\Delta t,x)+\bar{b_2}((k+1)\Delta t,x)\alpha$.

Take now S_k, $\tilde S_k\in\mathbb{R}^2$. By (99), the previously shown uniform continuity properties and uniform integrability in S^1_{k+1} of $p_\theta(S^1_{k+1}|S^1_k,S^2_k)$ and $\bar p(y_{k+1}|S^1_{k+1},\tilde S^1_k,\tilde S^2_k,\theta)$ we have that, given $\eta>0$, one can choose $\delta>0$ such that for $\|S_k-\tilde S_k\|<\delta$ it results that

$$\begin{aligned}
&|\phi_F(S_k,\theta,y_{k+1})-\phi_F(\tilde S_k,\theta,y_{k+1}|\\
\le&\int|\bar F(S^1_{k+1},S^1_k,S^2_k)\bar p(y_{k+1}|S^1_{k+1},S^1_k,S^2_k,\theta)\\
&\quad-\bar F(S^1_{k+1},\tilde S^1_k,\tilde S^2_k)\bar p(y_{k+1}|S^1_{k+1},\tilde S^1_k,\tilde S^2_k,\theta)|p_\theta(S^1_{k+1}|S^1_k,S^2_k)dS^1_{k+1}\\
&\quad+\int\bar F(S^1_{k+1},\tilde S^1_k,\tilde S^2_k)|p_\theta(S^1_{k+1}|S^1_k,S^2_k)-p_\theta(S^1_{k+1}|\tilde S^1_k,\tilde S^2_k)|\bar p(y_{k+1}|S^1_{k+1},\tilde S^1_k,\tilde S^2_k,\theta)dS_{k+1}\\
\le&\eta\int p_\theta(S^1_{k+1}|S^1_k,S^2_k)dS^1_{k+1}+\frac{\eta}{\gamma}\int\bar p(y_{k+1}|S^1_{k+1},\tilde S^1_k,\tilde S^2_k,\theta)dS^1_{k+1}=2\eta.
\end{aligned}\tag{102}$$

Thus showing the uniform continuity in S of $\phi_F(S,\theta,y)$ for each pair (θ,y) and with it also the induction step.

■

9 Appendix 3:

We derive the stochastic differential equation followed by $\ln(l_x(t))$ and hence obtain its volatility function.

From equation (9) and (13) we have

$$\ln l_x(t)=\bar b_0(t,x)+\sum_{i=1}^2\bar b_i(t,x)r(t,x_i)-\ln x.$$

Then

$$
\begin{aligned}
d\ln(l_x(t)) &= d[\bar{b}_0(t,x) + \sum_{i=1}^{2} \bar{b}_i(t,x) r(t,x_i)] \\
&= \bar{b}_0'(t,x)dt + \sum_{i=1}^{2} \bar{b}_i'(t,x) r(t,x)dt + \bar{b}_i(t,x) dr(t,x_i).
\end{aligned}
$$

Substituting the dynamics for $r(t,x_k)$ in (14) and simplying gives

$$
\begin{aligned}
d\ln(l_x(t)) = \Bigg[\bar{b}_0'(t,x) + \sum_{i=1}^{2} \bar{b}_i'(t,x) r(t,x_i) + \sum_{i=1}^{2} \bar{b}_i(t,x) \Bigg[\bar{b}_0'(t,x_i) + \sum_{k=1}^{2} \bar{b}_k'(t,x_i) r(t,x_k) \\
+\ \sigma(t,t+x_i)\bar{\sigma}(t,t+x_i)]]\, dt + \sum_{i=1}^{2} \bar{b}_i(t,x) e^{-\lambda x_i} g dw.
\end{aligned}
$$

Therefore

$$
var(d\ln(l_x(t)) = \left(\sum_{i=1}^{2} \bar{b}_i(t,x) e^{-\lambda x_i}\right)^2 g^2 \Delta t.
$$

10 References

BABBS, S.H. and NOWMAN K.B. (1999) : Kalman filtering of generalized Vasicek term structure models. *Journal of Financial and Quantitative Analysis.* 34, 115-130.

BHAR, R., CHIARELLA, C. and RUNGGALDIER, W. (2002): Estimation in models of the instantaneous short term interest rate by use of a dynamic Bayesian algorithm in *Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann, Eds. K. Sandmann and P. Schönbucher, Springer-Verlag*, 177-195.

BHAR, R., CHIARELLA, C. and TÔ, T-D (2002): A maximum likelihood approach to the estimation of Heath-Jarrow-Morton Models. Working Paper 80, Quantitative Finance Research Centre, UTS.

BØLVIKEN, E. and STORVIK, G. (2001) : Deterministic and stochastic particle filters in state-space models in *Statistics for Engineering and Information Science: Sequential Monte Carlo Methods in Practice, Eds. A. Doucet, N. de Freitas and N. Gordon, Springer-Verlag,* 97-116.

BRACE, A. and MUSIELA, M. (1994): A multifactor Gauss Markov implementation of Heath, Jarrow, and Morton. *Mathematical Finance* 3, 135-155.

CHAN, K., KAROLYI, A., LONGSTAFF, F. and SANDERS, A. (1992): An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance* 47, 1209-1227.

CHAPMAN, D., LONG, J. and PEARSON, N.(1999): Using proxies for the short rate: when are three months like an instant? *Review of Financial Studies* 12(4), 763-806.

CHIARELLA, C., HUNG, H. and TÔ, T-D (2005): The volatility structure of the fixed income market under the HJM Framework: A nonlinear filtering approach. Working Paper 151, Quantitative Finance Research Centre, UTS

CHIARELLA, C. and KWON, O.K. (2003): Finite dimensional affine realisations of HJM models in terms of forward rates and yields. *Review of Derivatives Research* 6(2), 129-155.

CHIARELLA, C., PASQUALI, S. and RUNGGALDIER, W.J. (2001) : On filtering in Markovian term structure models (an approximation approach), *Advances in Applied Probability*, 33(4), 794-809, 2001.

CHIB, S., NARDARI, F. and SHEPHARD, N. (2002) : Markov chains Monte Carlo methods for stochastic volatility models. *Journal of Econometrics* 108, 281-316.

EPISCOPOS A. (2000): Further evidence on alternative continuous time models of the short-term interest rate. *Journal of International Financial Markets, Institutions & Money* 10, 199-212.

HEATH, D., JARROW, R. and MORTON, A. (1992): Bond pricing and the term structure of interest rates: A new methodology for continent claim valuations. *Econometrica*, 60 (1), 77-105.

NOWMAN, K.B. (1997): Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance* 52, 1695-1706.

NOWMAN, K.B. (1998): Continuous-time short term interest rate models. *Applied Financial Economics* 8, 401-407.

NOWMAN, K.B. (2001): Gaussian Estimation and forecasting of multi-factor term structure models with an application to Japan and the United Kingdom. *Asia Pacific Financial Markets* 8, 23-34.

NOWMAN, K.B. (2003) : A note on Gaussian estimation of the CKLS and CIR models with feedback effects for Japan. *Asia Pacific Financial Markets* 10, 275-279.

POLSON, N. and Tew, B. (2000) : Bayesian portfolio selection: an empirical analysis of the S&P index 1970-1996. *Journal of Business and Economic Statistics* 18, 164-173.

SUN, L. (2003): Nonlinear drift and stochastic volatility: An empirical investigation of short-term interest rate models. *The Journal of Financial Research* XXVI(3), 389-404.